



# Étalonnage automatique d'un système d'acquisition Caméras - Centrale inertielle - Lidar 3D

Clement Deymier, Céline Teulière, Thierry Chateau

## ► To cite this version:

Clement Deymier, Céline Teulière, Thierry Chateau. Étalonnage automatique d'un système d'acquisition Caméras - Centrale inertielle - Lidar 3D. Reconnaissance de Formes et Intelligence Artificielle (RFIA) 2014, Jun 2014, France. hal-00988741

**HAL Id: hal-00988741**

**<https://hal.science/hal-00988741>**

Submitted on 9 May 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Étalonnage automatique d'un système d'acquisition Caméras - Centrale inertielle - Lidar 3D

C. Deymier<sup>1</sup>

C. Teulière<sup>1</sup>

T. Chateau<sup>1</sup>

<sup>1</sup> Institut Pascal UMR 6602 CNRS-UBP-IFMA

clement.deymier@univ-bpclermont.fr

## Résumé

*Cet article présente une méthode entièrement automatique d'étalonnage de systèmes d'acquisition complexes comprenant une ou plusieurs caméras, une centrale inertielle et un lidar 3D. Le principe consiste à estimer les paramètres intrinsèques et extrinsèques en mettant en correspondance des primitives détectées dans les images des caméras avec le nuage de points 3D fourni par le télémètre. Ce travail propose une formalisation mathématique permettant l'unification des trois types de capteurs au sein d'une même fonction de vraisemblance, une stratégie pour l'évaluation rapide de contraintes entre les données images et les données télémétriques, et enfin, l'utilisation d'un algorithme de minimisation à trois familles de paramètres qui permet une estimation simultanée de tous les paramètres d'étalonnage. Des expériences réalisées sur des systèmes d'acquisition synthétiques permettent d'évaluer le domaine de convergence de l'approche proposée ainsi que ses performances en termes de précision et de robustesse en présence de bruit.*

## Mots Clef

Étalonnage automatique, Optimisation, Ajustement de faisceau, Centrale inertielle, Télémètre 3D.

## Abstract

*This article presents a fully automated calibration method suitable for complex acquisition systems comprising one or more cameras, an inertial measurement unit and a 3D lidar. The principle is to estimate the intrinsic and extrinsic parameters by matching features detected in the camera images with the 3D point cloud provided by the rangefinder. This work proposes a mathematical formalization for the unification of the three types of sensors within the same likelihood function and a minimization algorithm with three families of parameters that enables the simultaneous estimation of all calibration parameters. Experiments conducted on synthetic acquisition systems are presented and assess the area of convergence of the proposed approach and its performance in terms of accuracy, correctness and robustness in the presence of noise.*

## Keywords

Automatic calibration, optimisation, Bundle-adjustment, IMU, 3D Lidar.

## 1 Introduction

Dans un but de cartographie, l'exploitation des données télémétriques nécessite de recalcr les données fournies dans un même référentiel. En effet, les senseurs télémétriques mesurent uniquement une distance et une direction dans leur propre référentiel donc, lors du déplacement du système d'acquisition, le référentiel du télémètre se déplace aussi par rapport à l'environnement. Afin de construire un modèle tridimensionnel d'une scène, il est donc impératif de connaître d'une part le déplacement grâce à un organe de localisation précis et d'autre part la transformation géométrique permettant de passer du référentiel de la localisation vers le référentiel du télémètre : ce sont les paramètres extrinsèques. Identiquement, l'utilisation de caméras nécessite elle aussi une localisation afin de déterminer les poses (orientation et translation) auxquelles les images sont acquises. La connaissance de cette information permet de recalcr les images avec le nuage de points 3D obtenu par le senseur télémétrique et effectuer des traitements, par exemple : la coloration du nuage de points, la détection des objets mobiles... Encore une fois, un organe de localisation peut fournir ces informations tant que l'on dispose d'un étalonnage extrinsèque et intrinsèque des caméras.

Les centrales inertielles de haute qualité permettent d'effectuer ce recalage grâce à leur grande vitesse de rafraîchissement ainsi que leur précision. Leur principe repose sur l'utilisation simultanée de deux organes de mesure différents : (i) un accéléromètre qui mesure l'accélération du système et (ii) un gyromètre mesurant les vitesses angulaires. Les mesures sont ensuite intégrées deux fois afin de restituer les poses 3D de la centrale inertielle. Ces systèmes appartiennent à la catégorie des senseurs proprioceptifs et, comme l'odométrie, ils dérivent au cours du temps à cause de biais et d'incertitude de mesure. De ce fait, ces outils sont communément couplés à un système de localisation extéroceptif, par exemple, un GPS centimétrique.

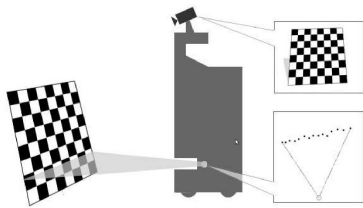


FIGURE 1 – Illustration du processus d’étalonnage caméra-télémètre utilisant une mire plane. Le système d’acquisition, ici un robot, perçoit par son télémètre un plan et la caméra détecte le damier. Les deux informations sont intégrées au sein d’une fonction objectif tandis qu’un algorithme de minimisation retrouve le paramétrage extrinsèque du système.

Dans cet article, nous allons présenter une méthode d’étalonnage permettant l’estimation des paramètres extrinsèques (les transformations géométriques entre les capteurs) et l’étalonnage intrinsèque des capteurs d’un système d’acquisition comportant une centrale inertielle, un capteur télémétrique et une ou plusieurs caméras. Premièrement, nous parcourons les différentes approches de la littérature dans la section suivante, deuxièmement, nous formaliserons cette problématique dans la section 2 sous la forme d’une fonction de vraisemblance, dont la minimisation à l’aide d’un algorithme d’optimisation 3.3 permet l’obtention des valeurs d’étalonnage. Les résultats obtenus seront analysés en termes de précision et de robustesse dans la section expérimentation 4.

## 1.1 Travaux antérieurs

La littérature présente déjà de nombreuses solutions d’étalonnage entre un télémètre et une caméra. La plupart des travaux se basent sur l’observation conjointe d’une mire plane possédant un motif damier par les deux capteurs [Mei, 2007, Florez et al., 2008, Vasconcelos et al., 2012]. L’approche consiste à détecter le plan, puis, après une formulation théorique des équations de projection, d’effectuer la minimisation d’un critère (figure 1) visant à contraindre les mesures télémétriques sur la surface plane. Ces solutions contraignantes nécessitent que les capteurs soient statiques pendant l’étalonnage. Il est donc impossible d’étalonner à posteriori une séquence déjà acquise. Des techniques plus automatiques ont vu le jour, l’une d’elles se base sur l’information mutuelle [Pandey et al., 2012] entre l’image et l’intensité de retour d’un télémètre laser. Il existe en effet une corrélation entre les variations de texture de l’image d’une caméra et la variation d’intensité de retour du laser. La maximisation de cette quantité d’information permet de remonter aux paramètres extrinsèques liant télémètre et caméra sans la nécessité de présenter une mire au système d’acquisition.

L’étalonnage entre un système de caméra et une centrale inertielle est un problème peu traité. Un article pu-

blié [Lobo and Dias, 2007] présente une méthode très intéressante, car elle permet de retrouver l’étalonnage extrinsèque entre ces deux capteurs et ainsi profiter de leurs avantages respectifs en fusionnant les informations. Cette solution est basée sur l’observation d’une mire verticale permettant de localiser la caméra tandis que des équations de changement de repère unifient les données d’accélération et les données caméra. Au final, l’algorithme fournit la transformation affine estimée. Cette solution a le désavantage de n’avoir rien d’automatique. De plus, elle nécessite une mire et reste spécifique à cette configuration.

L’étalonnage entre un capteur télémétrique et une centrale inertielle a aussi été effectué en utilisant l’algorithme ICP (Iterative Closest Point) pour recaler les données télémétriques et trouver les paramètres extrinsèques liant la trajectoire de la centrale inertielle et celle du télémètre. L’article [Zhu and Liu, ] estime ces mêmes paramètres grâce à l’extraction d’un plan parmi les mesures télémétriques constituant le sol supposé plan. En plus de faire l’hypothèse de l’observation de planéité du sol, le principe rend impossible l’observation d’un angle et de deux translations de la transformation. L’article contourne cette difficulté en détectant des poteaux dans la scène.

Aucune des approches citées ne permet une unification réelle de tous les capteurs au sein d’un même espace de travail de manière automatique et sans l’utilisation d’une mire ou d’une configuration spécifique. L’objectif de nos travaux est de développer une méthode d’étalonnage caméra-système inertiel-télémètre flexible et ne posant aucun a priori sur la scène observée. La méthode que nous présentons ici est de plus capable de traiter de manière égale les données d’un capteur télémétrique tournant quelle que soit sa configuration d’acquisition : qu’il observe plusieurs fois la scène ou qu’il ne l’observe qu’une seule fois comme les télémètres rotatifs perpendiculaires au déplacement du système d’acquisition.

## 2 Formalisation

L’objectif de notre approche est d’estimer :  $T_{imu}^t$  la transformation géométrique (*ie* une matrice transformation homogène) permettant le passage de la centrale inertielle au repère du télémètre,  $T_{imu}^{c_j}$  les transformations de la centrale inertielle aux caméras  $\mathcal{C} = \{c_j\}$  ainsi que les paramètres intrinsèques des caméras  $\theta_j$  (voir figure 2). La composition  $T_{imu}^{c_j} T_{imu}^t{}^{-1}$  permet de définir une transformation entre la caméra et le capteur télémétrique directement. Afin d’obtenir une méthode fonctionnant sans l’aide d’une mire, d’une configuration particulière ou d’une trajectoire singulière, nous allons obligatoirement devoir comparer les informations caméra *avec* les informations télémétriques. Pour cela, nous nous sommes orientés vers une approche basée primitive consistant à détecter des points d’intérêts dans les images puis à les trianguler dans l’espace pour finalement les comparer avec le nuage de mesures télémétriques 3D afin d’en extraire un étalonnage. On suppose dans ce travail disposer de données télémétriques denses et

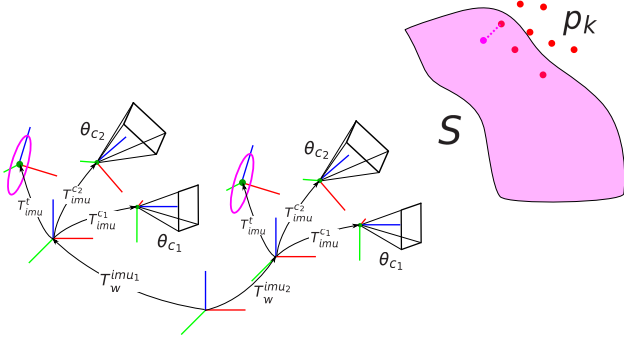


FIGURE 2 – Cette figure illustre un système d'acquisition à deux instants différents. Soit  $T_w^{imu1}$  et  $T_w^{imu2}$  les poses du système inertiel à ces deux moments.  $T_{imu}^{c1}$  et  $T_{imu}^{c2}$  sont les paramètres extrinsèques à retrouver et transformant la pose de la centrale inertielle vers le repère des caméras.  $T_{imu}^t$  représente l'étalonnage extrinsèque du télémètre.  $\theta_{c1}$  et  $\theta_{c2}$  le paramétrage intrinsèque des caméras.  $\mathcal{P}$  l'ensemble des points 3D issu de la géométrie reconstruite grâce aux images et  $S$  la surface engendrée par les données télémétriques. L'objectif est de minimiser les erreurs de reprojection des  $p_k$  dans les images ainsi que l'erreur entre  $p_k$  et la surface  $S$  afin de restituer toutes les valeurs d'étalonnage.

d'une connaissance a priori des valeurs approximatives des paramètres d'étalonnage. De plus, nous considérons ici que les poses de la centrale inertielle notées  $T_{imu_i}$  sont localement infiniment précises. Ce qui se justifie dans la réalité lors de l'utilisation d'une centrale inertielle de haute qualité.

Soit  $F_{(k,u,r)}$  le  $r$ -ième point d'intérêt détecté au sein de l'image d'index  $k$  provenant de la caméra d'index  $u$ . Soit  $\mathcal{P}$  l'ensemble des points 3D  $p$  de l'espace correspondant à la triangulation des observations du repère image. On peut alors exprimer la vraisemblance  $L_1$  des observations caméra en fonction des paramètres  $\theta_j$ ,  $T_{imu}^{c_j}$  et des poses de la centrale inertielle notées  $T_{imu_i}$  sous la forme :

$$L_1 = \prod_{k=1}^n \prod_{u=1}^m \prod_{r=1}^l f(F_{(k,u,r)}; \theta_u, T_{imu}^{c_u}, T_{imu_k}), \quad (1)$$

où  $f$  est une densité de probabilité,  $n$  est le nombre d'images de la séquence,  $m$  le nombre de caméras du système multi-caméra,  $l$  le nombre de primitives présentes dans l'image et  $\mathcal{F}$  l'ensemble des primitives présentes dans la séquence. Sous l'hypothèse que le bruit du détecteur de primitives de l'image a un comportement gaussien d'écart type  $\sigma_g$  on obtient :

$$L_1 = \prod_{k=1}^n \prod_{u=1}^m \prod_{r=1}^l \exp \left( - \frac{(\pi(\theta_u, T_{imu}^{c_u}, T_{imu_k}, p(k, u, r)) - F_{(k,u,r)})^2}{2\sigma_g^2} \right), \quad (2)$$

avec  $\pi$  la fonction de projection dans la caméra  $u$  dépendante des paramètres  $\theta_u$  et  $p(k, u, r)$  le point 3D associé à l'observation  $F_{(k,u,r)}$ . En composant par la fonction  $g(x) = -\log(x)$ , nous obtenons :

$$-\log(L_1) = \frac{1}{2\sigma_g^2} \sum_{k=1}^n \sum_{u=1}^m \sum_{r=1}^l (\pi(\theta_u, T_{imu}^{c_u}, T_{imu_k}, p(k, u, r)) - F_{(k,u,r)})^2 \quad (3)$$

Pour l'instant, les données télémétriques n'ont pas été insérées dans la formalisation. Nous constatons que la maximisation de la vraisemblance correspond dans le cadre gaussien à une minimisation du carré des erreurs. Cette erreur est nommée "erreur de reprojection", car elle correspond à la distance entre la projection d'une primitive 3D et son observation dans le repère image. La minimisation de cette fonction et l'obtention des paramètres optimaux seront abordées dans la section 3.3 et font intervenir des algorithmes d'optimisation non linéaire.

L'intégration des données télémétriques s'effectue de la même manière. Les données télémétriques définissent une surface sur laquelle devraient se trouver les primitives issues des images. En effet, les primitives correspondant à des objets de la scène, elles sont localisées à la surface de ces objets. Nous allons donc contraindre par l'ajout d'un critère supplémentaire les primitives 3D issues des images à être présentes sur la surface définie par les données télémétriques. En reprenant l'équation (1) et en admettant que la surface possède un bruit gaussien  $\sigma_s$  on obtient le nouveau critère :

$$-\log(L_2) = -\log(L_1) + \sum_{v=1}^q \frac{(\|p(v) - S(T_{imu}^t, T_{imu})\|)^2}{2\sigma_s^2} \quad (4)$$

où  $S(T_{imu}^t, T_{imu})$  représente la surface engendrée par les mesures télémétrique. Cette surface est dépendante des poses de la centrale inertielle  $T_{imu}$ , mais aussi de la transformation entre la centrale inertielle et le senseur télémétrique  $T_{imu}^t$  qui est le paramètre à estimer. Il est donc nécessaire de remettre à jour la surface à chaque fois qu'une modification de ces paramètres est effectuée. Nous allons voir dans la section suivante comment estimer de manière cohérente la distance entre la primitive et la surface  $\|p(v) - S(T_{imu}^t, T_{imu_s})\|$ .

## 2.1 Distance à la surface engendrée

La surface  $S$  n'est pas directement observable. En réalité, nous ne la connaissons que par l'intermédiaire des points de mesures télémétriques. Afin de définir totalement le score d'optimisation exprimé par l'équation (4), il est nécessaire d'estimer une distance entre la primitive 3D et la surface déterminée par les données télémétriques.

À ces fins, nous avons choisi d'utiliser la distance du point  $p$  au point de la surface  $S$  le plus proche. Cependant, dans le cas où la surface est échantillonnée de manière dense, nous pouvons remplacer la distance à la surface par la distance à la mesure télémétrique la plus proche, ce qui se formalise par :  $\|p - S(T_{imu}^t, T_{imu_s})\| = \min_{\mathcal{Z}} \|p - z_i\|$  dans

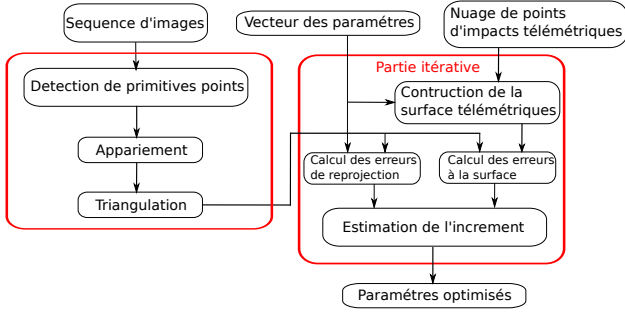


FIGURE 3 – Principe général de la méthode d'étalonnage.

lequel  $\mathcal{Z} \doteq \{z_k\}_{k=1,\dots,n}$  représente l'ensemble des points d'impact télémétriques exprimés dans le repère monde.

### 3 Principe général de la méthode

Le problème dont la formulation théorique a été présentée dans la partie précédente est résolu en minimisant la fonction objectif intégrant les différents paramètres intrinsèques et extrinsèques. Cette méthode se déroule en 5 étapes successives. Premièrement, des primitives de type point sont extraites des images grâce à un détecteur présenté dans la section 3.2. Dans un second temps, des descripteurs locaux de texture sont extraits autour de chaque point afin de pouvoir les identifier et les comparer aux autres. Les primitives sont associées entre les différentes images ce qui permet de déterminer la position tridimensionnelle du point par triangulation.

Une fois ces trois phases réalisées, un algorithme de minimisation non linéaire permet d'optimiser tous les paramètres à estimer. Celui-ci se déroule en deux parties majeures qui sont répétées de manière itérative :

(i) On effectue l'estimation des erreurs (aussi appelées résidus) qui sont composées des erreurs vision, calculées en reprojectant les primitives dans les images ainsi que des erreurs à la surface télémétrique, calculées grâce à la distance entre les primitives et la surface comme défini à la section 2.1. (ii) Un incrément est estimé et corrige le vecteur de paramètres, étape présentée en détail dans la section 3.3. La minimisation s'arrête lorsque les erreurs de reprojection et les erreurs de distance entre les points triangulés et la surface télémétrique sont minimales. Une vision schématique du processus est représentée par la figure 3.

#### 3.1 Modélisation du système

Les caméras sont modélisées à l'aide du modèle unifié [Barreto, 2006]. Il s'est avéré être capable de modéliser avec le même formalisme mathématique autant les caméras perspectives que les caméras grand-angles de type Fish-Eye. Contrairement au modèle sténopé dont l'étape de projection s'effectue directement sur le plan focal, le modèle unifié passe par une projection sur la sphère puis une seconde phase de projection sur un plan. Cette solution originale a la propriété d'être une excellente modélisation des phénomènes physique donnant lieu à la distorsion géométrique.

trique.

La paramétrisation extrinsèque, c'est-à-dire les transformations géométriques sont représentées par trois paramètres de translation exprimés par leurs coordonnées cartésiennes et trois paramètres de rotation. Ceux-ci sont exprimés grâce à un atlas de carte exponentielle locale comme expliqué dans les travaux de [Taylor and Kriegman, 1994].

#### 3.2 Détection d'amer, appariement et triangulation

Les primitives de type points sont détectées en utilisant le détecteur de point d'intérêt de Harris et Stephen [Harris and Stephens, 1988]. L'appariement de celles-ci est réalisé à l'aide de descripteurs locaux de type *ZNCC* permettant l'invariance à la luminosité de la scène identiquement aux travaux de [Royer et al., 2007].

Puisque des paramètres extrinsèques et intrinsèques approximatifs sont supposés connus, nous connaissons approximativement la pose de chaque image. La restitution en 3D des primitives points s'effectue donc directement par triangulation à partir des points d'intérêt à l'aide de la méthode du point du milieu adapté au modèle unifié.

#### 3.3 Minimisation

La minimisation de la fonction de vraisemblance est effectuée à l'aide d'un algorithme d'optimisation non-linéaire. La problématique est donc l'optimisation conjointe de l'ensemble des paramètres extrinsèques, des paramètres intrinsèques des caméras ainsi que la position de tout les points 3D triangulés. Ce cas d'utilisation est très proche de la notion d'ajustement de faisceaux classiquement utilisé en "Structure from Motion". La seule différence réside dans le fait que les poses des images ne sont pas optimisées. Pour cette raison, nous avons utilisé un algorithme classique : Levenberg-Marquardt [Marquardt, 1963] adapté pour l'optimisation de trois famille de paramètres. Ceci se formalise par un pas et une direction d'optimisation définis par :

$$X_{k+1} = X_k - (J^T J + \lambda I)^{-1} J^T \epsilon \quad (5)$$

Dans lequel  $J$  est la jacobienne du système,  $\lambda$  le coefficient d'amortissement,  $X$  le vecteur comprenant l'ensemble des paramètres à optimiser aux itérations  $k$  et  $k+1$ .  $\epsilon$  représente le vecteur des résidus.

**Estimation rapide des résidus.** Les résidus sont constitués de deux parties, l'une correspondant aux erreurs de reprojection des primitives et l'autre aux distances entre les primitives et la surface télémétrique. Les premières sont calculées de la manière suivante : pour chaque primitive point 3D  $p$  de la scène et chaque observation  $\nu$  fournie par le détecteur de Harris dans une image, la primitive est projetée dans le repère image grâce à la fonction de projection du modèle unifié. Le vecteur des résidus pour cette primitive dans cette image est défini par  $\vec{e}(p, \nu) = \pi(p, X) - \nu$ , où  $X$  est le vecteur de paramètres de l'optimisation contenant les paramètres extrinsèques et intrinsèques de la caméra. Pour les résidus liés à la distance entre la primitive

3D et la surface télémétrique, il est nécessaire de régénérer la surface télémétrique pour la valeur des paramètres extrinsèques du télémètre fournie par l'optimisation. À ces fins, nous effectuons un changement de repère afin d'exprimer les données télémétriques dans le repère local au capteur dans un repère global à l'aide de la centrale inertielle. Cette transformation se formalise par  $T_{imu_i}^t = T_{imu}^t T_{imu_i}$  où  $T_{imu}^t$  représente les paramètres extrinsèques du télémètre (présent dans  $X$ ) et  $T_{imu_i}$  la  $i$ -ème pose de la centrale inertielle correspondant à la même date que la donnée télémétrique à recalculer. La fréquence d'acquisition du senseur télémétrique étant souvent supérieure à la fréquence de rafraîchissement de la centrale inertielle, nous interpolons linéairement ces transformations à la date de chaque point télémétrique. Nous avons choisi dans la section 2.1 d'utiliser la distance minimale entre les primitives et la surface comme résidu. À ces fins, une recherche de voisinage est employée pour trouver la donnée télémétrique la plus proche de chaque primitive. Cependant, face au volume très important de mesures télémétriques et au grand nombre de primitives détecté par image, nous exploitons un arbre de partitionnement pour accélérer ces requêtes. Les points d'impact des mesures télémétriques sont insérés dans un Octree à chaque itération de l'algorithme d'optimisation, et pour chaque primitive, nous recherchons le point  $q$  le plus proche grâce à un algorithme de voisinage classique. Le résidu fourni à la minimisation est :  $\bar{\epsilon}(p, q) = p - q$

**Évaluation des écart-types des mesures.** La formulation de la fonction de coût (4) fait intervenir les écart-types des bruits de mesures,  $\sigma_g$  pour le bruit des détections dans le repère image et  $\sigma_s$  pour l'écart type relatif au bruit de la surface télémétrique utilisée pour le calcul des résidus.  $\sigma_s$  représentant le bruit de la surface, il est donc choisi égal à l'écart type de mesure du capteur télémétrique. Cette donnée est disponible dans les documents techniques joints au senseur.  $\sigma_g$  représentant l'écart-type sur la détection des points de Harris, nous employons la valeur  $2px$  communément utilisée comme approximation par la communauté scientifique. Notons que la matrice de covariance du détecteur de Harris peut être calculée de manière analytique par  $\Sigma = \sigma_p^2 H(\nu)^{-1}$  où  $\sigma_p$  représente le bruit photogrammétrique pixellique de la caméra,  $H(\nu)$  la matrice de Harris estimée au coin détecté  $\nu$ .

## 4 Expérimentation

### 4.1 Méthodologie

Afin de disposer d'une vérité terrain de l'étalonnage d'un système d'acquisition, nous avons décidé d'utiliser un simulateur de capteur réaliste (4D-Virtualiz [Malarte, 2011]) permettant la génération de données avec un bruit réaliste. Dans le cadre de cet article nous considérerons une trajectoire d'étalonnage d'une centaine de mètres disposant de plusieurs virages (voir figure 4). Le système d'acquisition est composé d'une caméra et d'un

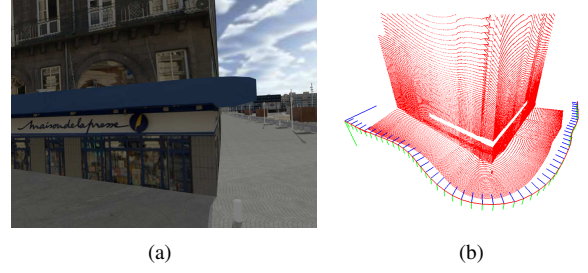


FIGURE 4 – La figure (a) représente une des images de l'environnement acquises par la caméra. (b) est une vue 3D des mesures télémétriques recalées (en rouge) et de la trajectoire du système d'acquisition (les repères tricolores).

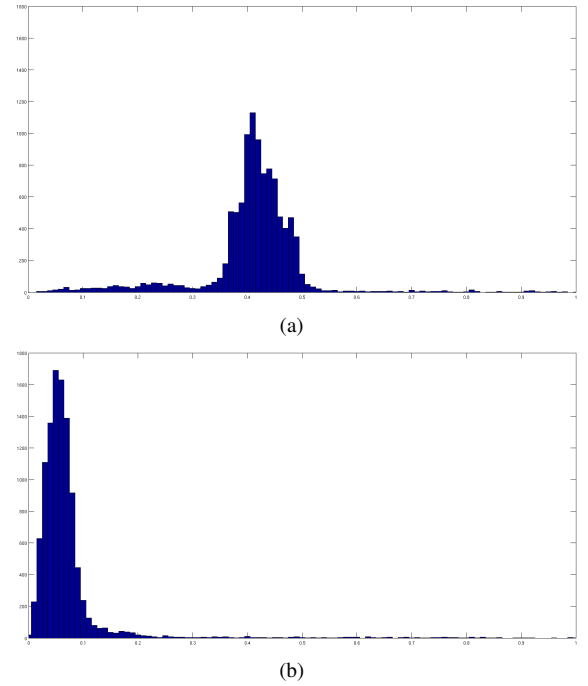


FIGURE 5 – Histogramme des résidus à la surface avant et après optimisation.

télémètre Riegl tout deux disposés sur le toit d'un véhicule virtuel et perpendiculairement à sa direction d'avancement. La caméra effectue une prise d'images couleur à une fréquence de  $10Hz$  tandis que le Riegl permet la collecte de plus de  $3.10^5$  point 3D par seconde, sa précision est centimétrique. Le véhicule virtuel est conduit manuellement à l'aide d'un joystick.

### 4.2 Resultats

Nous avons introduit des erreurs dans tous les paramètres extrinsèques et intrinsèques initiaux afin d'estimer la capacité de l'algorithme à retrouver les paramètres d'étalonnage du système. Dans l'exemple présenté à la figure 4, l'erreur de translation des paramètres extrinsèques est d'environ  $20cm$  pour la caméra et  $50cm$  pour le télémètre tandis que les orientations ont été entachées d'un biais de cinq degrés

selon tous les angles. De plus, le paramètre de distorsion du modèle intrinsèque de la caméra a été incrémenté de 0.1.

Les histogrammes 5 illustrent la répartition des résidus à la surface avant et après optimisation. Initialement centrés autour de la valeur 41cm, ils sont minimisés durant l'optimisation et se regroupent proche de zéro après convergence. Les écarts relatifs entre les paramètres extrinsèques estimés et la vérité terrain sont, en coordonnées cartésiennes, environ :  $\Delta_c = (0.0045m, 0.00014m, -0.0100m)$  pour la caméra,  $\Delta_t = (1.2053m, -0.0969m, 0.0951m)$  pour le télémètre, et enfin ( $\Delta f_x = 1.641, \Delta f_y = 0.655, \Delta u_0 = 1.097px, \Delta v_0 = 0.232px, \Delta \xi = 1.053e - 05$ ) pour l'étalonnage intrinsèques. Les erreurs angulaires ne dépassent pas le degré d'angle.

On remarque que notre approche a permis de retrouver avec précision les paramètres réels sauf pour le premier terme de translation du télémètre. En effet, la trajectoire utilisée étant planaire, elle ne permet pas de contraindre de manière suffisante les paramètres extrinsèques sur l'axe vertical. Le lecteur intéressé pourra se référer à [Lébraly, 2012] pour une analyse approfondie des cas singuliers de l'étalonnage des systèmes multi-capteurs. Le domaine de convergence est limité par deux étapes critiques. En effet, si les paramètres initiaux de la caméra ne sont pas assez précis, l'étape de triangulation échoue et l'optimisation ne peut avoir lieu. D'un autre côté, si les paramètres initiaux du télémètre sont erronés, la surface télémétrique est trop distante des points triangulés et l'optimisation échoue dans un minimum local, ou, l'ensemble des résidus est considéré comme aberrant. Pour cette séquence, l'amplitude des paramètres admissibles pour la translation du télémètre est de 70cm. Cela permet d'utiliser des valeurs recueillies à l'aide d'un mètre mesureur comme initialisation.

## 5 Conclusion

Nous avons présenté dans cet article une méthode d'étalonnage automatique d'un système d'acquisition possédant une ou plusieurs caméras, une centrale inertielle et un capteur télémétrique de type Lidar 3D. Nous avons défini une formalisation permettant l'unification de tous les paramètres d'étalonnage, extrinsèques et intrinsèques au sein d'une unique fonction de vraisemblance. La minimisation de celle-ci par un algorithme d'optimisation fournit une estimation précise des paramètres d'étalonnage. Les expérimentations ont montré que cette méthode est performante et permet, pour un grand domaine de convergence, de restituer avec exactitude les paramètres recherchés. Dans le cadre de ces travaux nous avons supposé disposer des paramètres approximatifs permettant la triangulation, l'appariement de primitive puis l'optimisation. Une extension possible consisterait à utiliser des méthodes de "Structure from motion" telles que la décomposition de la matrice fondamentale, l'ajustement de faisceau incrémental afin de les déterminer, agrandissant d'autant le cadre d'utilisation de cette méthode.

## Références

- [Barreto, 2006] Barreto, J. P. (2006). A unifying geometric representation for central projection systems. *Computer Vision and Image Understanding*, 103(3) :208–217.
- [Florez et al., 2008] Florez, S. A., Fremont, V., Bonnifait, P., et al. (2008). Influence of intrinsic parameters over extrinsic calibration between a multi-layer lidar and a camera. *IEEE Int. Conf. on Intelligent Robots and Systems*.
- [Harris and Stephens, 1988] Harris, C. and Stephens, M. (1988). A combined corner and edge detector. In *In Proc. of Fourth Alvey Vision Conference*, pages 147–151.
- [Lébraly, 2012] Lébraly, P. (2012). *Etalonnage de caméras à champs disjoints et reconstruction 3D : Application à un robot mobile*. PhD thesis, Université Blaise Pascal-Clermont-Ferrand II.
- [Lee and Moore, 2005] Lee, P. Y. and Moore, J. B. (2005). Gauss-newton-on-manifold for pose estimation. *Journal of industrial and management optimization*, 1(4) :565.
- [Lobo and Dias, 2007] Lobo, J. and Dias, J. (2007). Relative pose calibration between visual and inertial sensors. *The International Journal of Robotics Research*, 26(6) :561–575.
- [Malartre, 2011] Malartre, F. (2011). *Perception intelligente pour la navigation rapide de robots mobiles en environnement naturel*. PhD thesis, Ecole Doctorale Science Pour l'Ingénieur de Clermont Ferrand.
- [Marquardt, 1963] Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics*, 11(2) :431–441.
- [Mei, 2007] Mei, C. (2007). *Couplage Vision Omnidirectionnelle et Télémétrie Laser pour la Navigation en Robotique/Laser-Augmented Omnidirectional Vision for 3D Localisation and Mapping*. PhD thesis, PhD thesis, INRIA Sophia-Antipolis, Project-team ARobAS.
- [Pandey et al., 2012] Pandey, G., McBride, J. R., Savarese, S., and Eustice, R. (2012). Automatic targetless extrinsic calibration of a 3d lidar and camera by maximizing mutual information. In *AAAI*.
- [Royer et al., 2007] Royer, E., Lhuillier, M., Dhome, M., and Lavest, J. M. (2007). Monocular vision for mobile robot localization and autonomous navigation. *Int. Journal of Computer Vision*, 74 :237–260. 10.1007/s11263-006-0023-y.
- [Taylor and Kriegman, 1994] Taylor, C. J. and Kriegman, D. J. (1994). Minimization on the lie group so (3) and related manifolds. *Yale University*.
- [Vasconcelos et al., 2012] Vasconcelos, F., Barreto, J., and Nunes, U. (2012). A minimal solution for the extrinsic calibration of a camera and a laser-range finder.
- [Zhu and Liu, ] Zhu, Z. and Liu, J. Unsupervised extrinsic parameters calibration for multi-beam lidars.